EXHIBIT V

TO DECLARATION OF S. MERRILL WEISS IN SUPPORT OF PLAINTIFF ACACIA MEDIA TECHNOLOGIES CORPORATION'S MEMORANDUM OF POINTS AND AUTHORITIES IN OPPOSITION TO ROUND 3 DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY UNDER 35 U.S.C. § 112 OF THE '992, '863, AND '702 PATENTS; AND SATELLITE DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY OF THE '992, '863, AND '720 PATENTS

1977 IEEE International Conference on Acoustics, Speech & Signal Processing

IEEE 1977 ICASSP

Held at the

Sheraton-Hartford Hotel
Hartford, Connecticut
May 9-11, 1977

Record

77CH1197-3 ASSP

APPLICATION OF QUADRATURE MIRROR FILTERS TO SPLIT BAND VOICE CODING SCHEMES

D. Esteban and C. Galand

IBM Laboratory
06610 La Gaude, France

Abstract

This paper deals with applications of Quadrature Mirror Filters (QMF) to coding of voice signal in sub-bands. Use of QMF's enables to avoid the aliasing effects due to samples decimation when signal is split into sub-bands. Each sub-band is then coded independently with use of Block Companded PCM (BCPCM) quantizers. Then a variable number of bits is allocated to each sub-band quantizer in order to take advantage of the relative perceptual effect of the quantizing error.

The paper is organized as follows :

- First, splitting in two sub-bands with QMF's is analysed.
- Then, a general description of a splitband voice coding scheme using QMF's is made.
- Finally, two coding schemes are considered, operating respectively at 16 KBps and 32 KBps. Averaged values of S/N performances are given when encoding both male and female voices. Comparisons are made with conventional BCPCM and CCITT A-Law.

Taped results will be played at the conference.

1) Introduction

Decomposition of the voice spectrum in sub-bands has been proposed by R. Crochiere et al. /1/ as a means to reduce the effect of quantizing noise due to coding. The main advantages of this approach are the following :

- first, to localize the quantizing noise in narrow frequency sub-bands, thus preventing noise interference between these sub-bands,
- second, to enable the attribution of bit resources to the various frequency bands according to perceptual criteria.

As a result, the quantizing noise is perceptually more acceptable, and the signal to noise ratio is improved.

The implementation proposed in /1/ is straightforward and takes advantage of a bank of non-overlapping band-pass filters. Unfortunately, for a non perception of aliasing effects due to decimation, this approach needs sophisticated band-pass filters. The split-band coding scheme we propose here avoids these inconveniences. Quasi perfect sub-band splitting can be achieved by use of Quadrature Mirror Filters (QMF) /2/ associated with decimation/interpolation techniques.

2) QMF band splitting

Principle

Let us consider for explanation purposes Fig. 1 in which we describe the decomposition of a sampled signal in two contiguous sub-bands, where :

- H_1 is a sampled half band low pass filter with an impulse response $h_1(n)$,
- H_2 is the corresponding half band mirror filter, i.e. which satisfies the following magnitude relation :

$$|H_1(e^{j\omega T})| = |H_2(e^{j(\frac{\omega s}{2}-\omega)T})| \qquad (1)$$

where $\omega s = 2\Pi f_s = 2\Pi/T$ denotes the sampling rate and $H_1(e^{j\omega T})$ denotes the Fourier Transform of $h_1(n)$.

- K_1 is a half band low pass filter with an impulse response $k_1(n)$ and K_2 is the corresponding mirror filter of K_1.

After frequency limiting to $f_s/2$, the signal $x(t)$ is sampled at f_s and filtered by H_1 and H_2. The obtained signals $x_1(n)$ and $x_2(n)$ represent respectively the low and high half-bands of $x(n)$. As their spectra occupy half the Nyquist bandwidth of the original signal, the sampling rate in each band can be halved by ignoring every second sample. For reconstruction, the signals $y_1(n)$ and $y_2(n)$ are interpolated by inserting one zero valued sample between each sample and filtered by K_1 and K_2 before being added to give the signal $s(n)$.

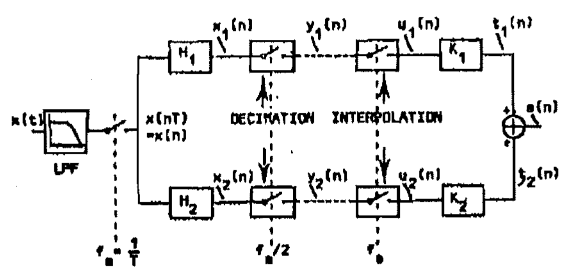

Fig 1 Principle of 2 sub-bands splitting by use of half-band mirror filters

Exhibit V Page 429

Let us analyse the structure of Fig. 1. If $X(z)$, $H(z)$ and $X_1(z)$ represent respectively the z transforms of $x(n)$, $h_1(n)$ and $x_1(n)$, then :

$$X_1(z) = H_1(z)X(z) \tag{2}$$

The z transform $Y_1(z)$ of the decimated signal $y_1(n)$ and the z transform $U_1(z)$ of the interpolated signal $u_1(n)$ are given by /3/ :

$$Y_1(z) = \frac{1}{2}(X_1(z^{\frac{1}{2}})+X_1(-z^{\frac{1}{2}})) \tag{3}$$

$$U_1(z) = Y_1(z^2) \tag{4}$$

After final filtering, the z transform of $t_1(n)$ is :

$$T_1(z) = K_1(z)U_1(z) \tag{5}$$

where $K_1(z)$ represents the z transform of $k_1(n)$.

Combining relations (2)-(5) gives :

$$T_1(z) = \frac{1}{2}(H_1(z)X(z)+H_1(-z)X(-z))K_1(z) \tag{6}$$

The z transform $T_2(z)$ is derived in a similar manner :

$$T_2(z) = \frac{1}{2}(H_2(z)X(z)+H_2(-z)X(-z))K_2(z) \tag{7}$$

The z transform $S(z)$ of the signal $s(n)$ is obtained by adding relations (6) and (7) :

$$\begin{aligned} S(z) &= \frac{1}{2}(H_1(z)K_1(z)+H_2(z)K_2(z))X(z) \\ &+\frac{1}{2}(H_1(-z)K_1(z)+H_2(-z)K_2(z))X(-z) \end{aligned} \tag{8}$$

The second term of this sum represents aliasing effects due to decimation and can be eliminated if we choose K_1 and K_2 appropriately. First, we must satisfy the symmetry relation (1). This is elegantly solved if H_1 is a finite impulse response (FIR) filter :

$$H_1(z) = \sum_{n=0}^{N-1} h_1(n)z^{-n} \tag{9}$$

It can be seen that the impulse response $h_2(n)$ of the mirror filter H_2 is obtained by inverting every second sample of $h_1(n)$.

$$H_2(z) = \sum_{n=0}^{N-1} h_1(n)(-1)^n z^{-n} = H_1(-z) \tag{10}$$

We can now cancel the second term of (8) by choosing :

$$K_1(z) = H_1(z) \tag{11}$$

$$K_2(z) = -H_2(z) = -H_1(-z) \tag{12}$$

Equation (8) now becomes :

$$S(z) = \frac{1}{2}(H_1^2(z)-H_1^2(-z))X(z) \tag{13}$$

Let us evaluate this relation on the unit circle

$$S(e^{j\omega T}) = \frac{1}{2}(H_1^2(e^{j\omega T}) - H_1^2(e^{j(\omega+\frac{\omega_s}{2})T})X(e^{j\omega T}) \tag{14}$$

If we choose for H_1 a symmetrical FIR filter, its Fourier transform $H_1(e^{j\omega T})$ can be expressed in term of its magnitude $H_1(\omega)$:

$$H_1(e^{j\omega T}) = H_1(\omega)e^{-j(N-1)\Pi\frac{\omega}{\omega_s}} \tag{15}$$

Substituting in (14) gives :

$$\begin{aligned} S(e^{j\omega T}) &= \frac{1}{2}(H_1^2(\omega)-H_1^2(\omega+\frac{\omega_s}{2})e^{-j(N-1)\Pi}) \\ &\times e^{-j(N-1)2\Pi\frac{\omega}{\omega_s}} \cdot X(e^{j\omega T}) \end{aligned} \tag{16}$$

Two cases are to be considered, depending on the parity of N .

. First case, N even

$$S(e^{j\omega T}) = \frac{1}{2}(H_1^2(\omega)+H_1^2(\omega+\frac{\omega_s}{2}))e^{-j(N-1)\omega T} . X(e^{j\omega T}) \tag{17}$$

Considering the case of perfect filters,

$$H_1^2(\omega) + H_1^2(\omega+\frac{\omega_s}{2}) = 1 \tag{18}$$

we get :

$$S(e^{j\omega T}) = \frac{1}{2} e^{-j(N-1)\omega T} . X(e^{j\omega T}) \tag{19}$$

or

$$s(n) = \frac{1}{2} x(n-N+1) \tag{20}$$

The signal is perfectly reconstructed (neglecting the gain factor 1/2) with a delay of (N-1) samples.

. Second case, N odd

In this case, the original signal cannot be perfectly reconstructed, it can be seen from (16) that the amplitude at $\omega = \omega_s/4$ is always zero.

To summarize, we have defined a set of conditions for perfect reconstruction :

H_1 = Symmetrical FIR filter of even order ;
$H_2(z) = H_1(-z)$;
$K_1(z) = H_1(z)$; $H_1^2(\omega) + H_1^2(\omega+\omega_s/2) = 1$;
$K_2(z) = -H_2(z)$;

Implementation

Fig. 2a gives an efficient implementation of the QMF band splitting, using a symmetrical FIR half band filter with an even number of coefficients. The input signal $x(t)$ is sampled at f_s and filtered by H_1 and H_2, giving the low-band channel $x_1(n)$ and the high-band channel $x_2(n)$. Then the sampling rate is decreased to $f_s/2$ by decimating every second sample, giving the signals $y_1(n)$ and $y_2(n)$.

Fig. 2b shows the reconstruction of the initial signal with the same filter. First, the sampling rate is increased to f_s by inserting one zero valued sample between each sample of $y_1(n)$ and $y_2(n)$, giving two signals $u_1(n)$ and $u_2(n)$. Then these signals are filtered by H_1 and H_2, and

the signal s(n) is obtained by subtracting the filtered signals $t_1(n)$ and $t_2(n)$.

The total number of multiplications to perform per initial sampling interval (splitting and reconstruction) is equal to the filter length N, the number of additions if of the order of N.

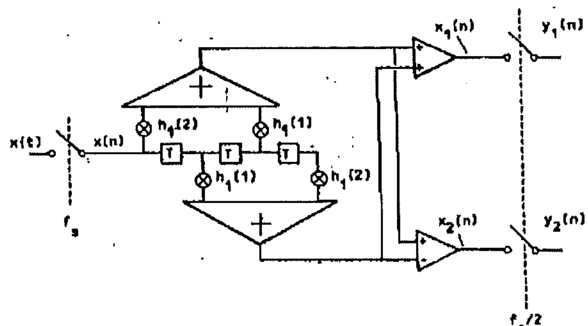

Fig 2a Quadrature channels splitting

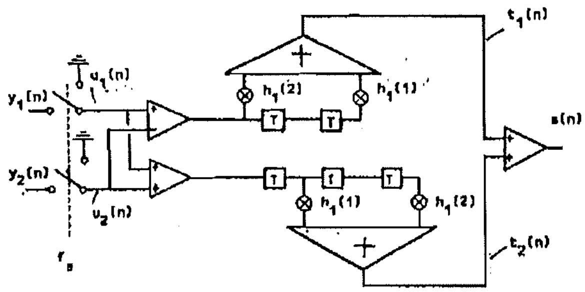

Fig 2b Quadrature channels reconstruction

3) Split-band voice coding scheme based on QMF 2^P sub-bands tree decomposition

In the previously described implementation, a signal x(t) was sampled at f_s to give a signal x(n), and split into two signals $y_1(n)$ and $y_2(n)$ with reduction of the sampling rate to $f_s/2$. This decomposition can be extended to more than two sub-bands by applying to $y_1(n)$ and $y_2(n)$, which represent respectively the low sub-band and the high sub-band of x(n), the same decomposition process as to the initial signal x(n) (see Fig. 4). Four signals are thus obtained with reduction of the sampling rate to $f_s/4$. The spectrum of each of these signals represents the spectrum of x(n) in the corresponding sub-band.

This decomposition can be generalized by repeating the processus p times. The initial signal is thus split into 2^P signals sampled at $f_s/2^P$ by a p-stage tree arrangement of decimation filters of the type shown on Fig. 2a. As the ith stage includes 2^{i-1} filters, the total number of filters is 2^{P-1}. The resulting information rate after p stages is the same as the one of the original signal.

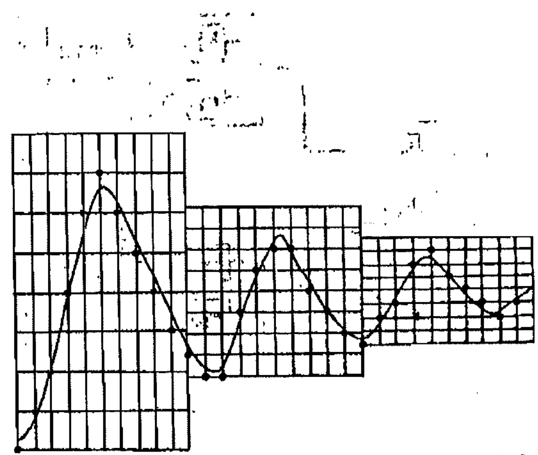

Fig 3 Block Companded PCM (BCPCM) principle

Quantization of the sub-band signals

As mentioned in /1/ and due to the fact that the sub-band signals are narrow band and Nyquist sampled, the sample-to-sample correlation of these signals is low. Consequently, straight PCM encoding techniques are preferred to differential methods.

An efficient and simple approach to code the sub-bands signals is obtained by means of Block Companded PCM (BCPCM) coding scheme /4/. This type of companding has been initially proposed for full band coding of speech waveforms, but can be straightforwardly applied to sub-band encoding. The principle of BCPCM coding can be summarized as follows :

- The samples are encoded on a block basis. For each block of M samples, a scale factor is chosen in such a way that the larger sample in the block will not fall out of the coded range.
- Then, the M samples of the block are quantized with respect to the obtained scale factor and both the coded values and the scale factor are transmitted.

The overhead bit rate necessary to the transmission of the scale factor is inversely proportional to the length of the blocks, but this length must be chosen so as to take in account the formant evolution. For a full band coding, a length of 8 to 16 ms has been found satisfying.

The main advantages of BCPCM are a low overhead information rate, a very large dynamic range, and no transient clipping. Fig. 3 shows the adaptation of the scale factor to the signal, considering three consecutive blocks, and assuming 3 bits quantization.

The BCPCM coding scheme has been used with success in conjunction with the QMF band splitting,

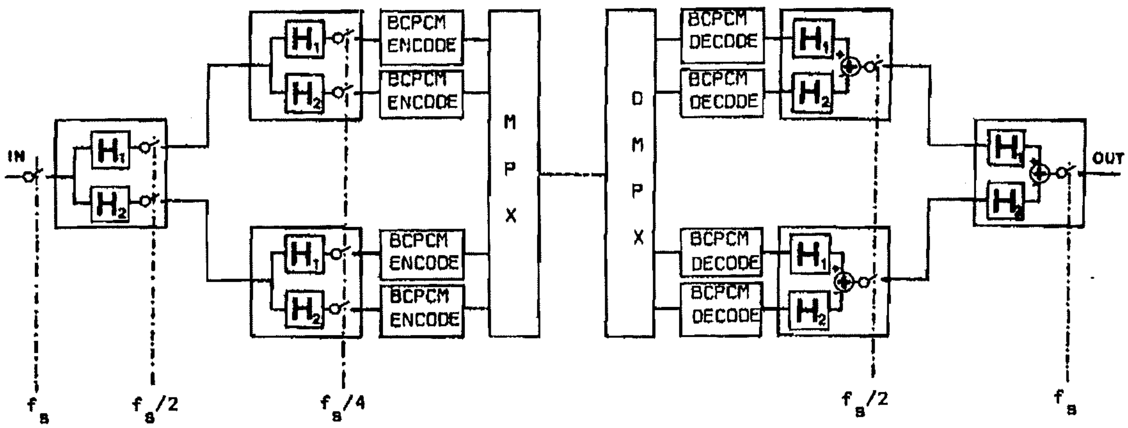

Fig 4 Four sub-bands SVCS with QMF and BCPCM

assuming different number of bits to code each frequency sub-band so as to weight the perceptual effect of the quantizing noise in the voice spectrum. Examples of bit allocation will be discussed in section 4. After quantization (see Fig. 4), the signals and scale factors from all channels are time multiplexed and transmitted.

2^P sub-bands reconstruction

At the receiving end, the data is demultiplexed and decoded. The reconstruction of the speech signal is made by a p-stage tree arrangement of filters of the type of the one shown in Fig. 2b.

If a same filter of N taps is used for each stage, the number of multiplies per input sample for the whole 2^P sub-bands decomposition/reconstruction is Np. In fact, filter constraint can be reduced from stage to stage with respect to the bandwidth so as to optimize the total processing. It has been shown in section 2, that there is a delay of (N-1) samples between the original and reconstructed signals in case of two sub-bands splitting. Consequently, the number of delayed samples is $(2^P-1)(N-1)$ for the 2^P sub-bands splitting.

4) Simulation of Split-band Voice Coding Scheme

In this section, two Split-band Voice Coding Schemes (SVCS) are considered. The first one operates at a bit rate of 16 KBps and provides a quality sufficient for telephony applications, the second operates at a bit rate of 32 KBps and gives a quality comparable to that provided by standard companded laws. The characteristics of these two coders are given hereafter.

- 16 KBps SVCS

  - input signal band limited to 0-4000 Hz
  - sampling rate : 8 KHz
  - number of sub-bands : 8
  - bit allocation : 3 3 3 1 1 1 1 1
  - block duration : 20 ms (160 samples)
  - number of overhead bits : 40

- 32 KBps SVCS

The characteristics of this coder are the same as the previous one, excepted the bit allocation that has been increased to :

5 5 5 4 4 3 3 1

- Performance

The performance of the two considered SVCS has been evaluated by comparison with conventional BCPCM coders operating at the same bit rate. For convenience, two types of BCPCM coders have been considered, the first one operating in PCM mode, the second one being able to take a PCM/DPCM decision /4/, so as to encode the high-correlated blocks of samples in differential mode.

The experimentations were made on a set of utterances pronounced by 7 speakers (4 female voices and 3 male voices) representing a total duration of 3.5 minutes of continuous speech. The averaged signal to noise ratios are given in table 1.

Table 1 Comparative performances (dB) of BCPCM and SVCS coders.

Coder \ Bit Rate	16 KBPS	32 KBPS
BCPCM (PCM Mode)	8	21
BCPCM (PCM/DPCM Mode)	11	24
SVCS	14	25

It must be noted that, for BCPCM coders, the PCM/DPCM decision enables a signal-to-noise improvement (SNRI) of 3dB. This improvement is not surprising and is in accordance with the well-known results of conventional PCM /5/. Moreover, it can be seen that split-band coding techniques provide SNRI over full-band techniques. This improvement is 3dB in case of 16 KBps, and only 1dB in case of 32 KBps. However, as noticed in /1/, it has been observed that for SVCS, the subjective level of the quantizing noise is less than for BCPCM, resulting in a more pleasant voice quality.

The previously described 16 KBps SVCS provides a speech quality which is sufficient for telephony applications. Furthermore, listening tests have shown that it is not possible to tell the difference between the 32 KBps SVCS and the CCITT 64 KBps A-Law, although the measured signal to noise ratios are respectively 25dB and 37dB.

5) Conclusions

The application of Quadrature Mirror Filters to Split-band Voice Coding Schemes has been discussed. As noticed in /1/, sub-band coding results in a signal to noise improvement over full-band coding. Moreover, the subjective effects of quantizing noise are less, resulting in a more pleasant coding quality.

Use of QMF enables to avoid aliasing effects due to decimation. Consequently, band splitting can be performed up to a large number of sub-bands without using sophisticated filters.

Two SVCS have been described, using BCPCM techniques and operating at 16 KBps and 32 KBps. The first one gives a speech quality which is sufficient for telephony applications. The second allows a quality comparable to that provided by the standard 64 KBps PCM code, thus achieving a halving of the bit rate for speech encoding.

References

/1/ R.E. Crochiere, S.A. Webber, J.L. Flanagan, "Digital coding of speech in sub-bands", 1976 Int'l IEEE Conf. on ASSP, Philadelphia.

/2/ A. Croisier, D. Esteban, C. Galand, "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques" 1976 Int'l Conf. on Information Sciences and Systems, Patras.

/3/ R. Schaffer, L. Rabiner, "A digital signal processing approach to interpolation", Proc. IEEE, Vol. 61, pp. 692-702, June 1973.

/4/ A. Croisier, "Progress in PCM and delta modulation : block companded coding of speech signals", 1974 Int'l Zürich seminar.

/5/ K.W. Cattermole, "Principles of pulse code modulation", London Iliffe Books Ltd.